DOCUMENT RESUME

ED 105 420 CS 001 747

AUTHOR Thomas, John W.
TITLE The Effects of Instruction on the Critical Thinking Abilities of Middle-School-Age Children.
PUB DATE Mar 75
NOTE 18p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 30-April 2, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Cognitive Processes; *Critical Thinking; *Educational Research; Elementary Secondary Education; *Instructional Materials; Middle Schools; Thought Processes

ABSTRACT

This study evaluated the effects of an experimental critical thinking program. Students at the sixth through ninth grade levels were randomly assigned to experimental and control groups. Experimental subjects received 36 two-hour weeks of instruction consisting of five independent units of programed booklets, games, and activities. Objective-referenced pretests and posttests were given to all subjects for each unit and a standardized critical thinking and dispositional inventory was administered at the end. Significant differences between treatments were found at all grade levels for units administered early in the study. However, results were mixed on later measures suggesting a debilitating motivational trend. (Author)

ED105420

504

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The Effects of Instruction on the Critical Thinking Abilities of Middle-School-Age Children

John W. Thomas
Humanizing Learning Program
Research For Better Schools, Inc.
Philadelphia, Pennsylvania

A paper presented at the Annual Meeting of the American Research Association, Washington, D.C., March, 1975.

CS001 742

The effects of instruction on the critical thinking abilities of middle-school age children

John W. Thomas

Research for Better Schools Inc.

Abstract

The present study evaluated the effects of an experimental critical-thinking program. Students at the sixth through ninth grade were randomly assigned to experimental and control groups. Experimental Ss received 36 two-hour weeks of instruction consisting of five independent units of programmed booklets, games, and activities. Objective-referenced pre- and posttests were given to all Ss for each unit and a standardized critical thinking and dispositional inventory was administered at the end. Significant differences between treatments were found at all grade levels for units administered early in the study, however, results were mixed on later measures suggesting a debilitating motivational trend.

The Effects of Instruction on the Critical Thinking Abilities of Middle-School-Age Children

John W. Thomas
Research for Better Schools, Inc.

Introduction

A number of studies have focused on the improvement of children's critical thinking abilities. Training studies have taken a variety of forms, including programs in inquiry, critical reading, formal logic, problem solving and reflective thinking. Curriculum development activities in critical thinking have taken place within all of the major subject matter areas. In his review of the processes of thinking, Russell (1965) summarized the diverse interpretations given to critical thinking as follows:

> *It has been made synonymous with the ability to abstract and organize information, to draw inferences, to search for relevant materials, to evaluate data, to compare sources, to employ a form-Missouri attitude, to distinguish fact from opinion, to detect propoganda, and to apply the rules of logical reasoning. (p. 14)*

Typically, the investigator interested in fostering critical thinking behaviors will define the objectives for an instructional program in terms of general abilities or aspects of critical thinking. These aspects may refer to activities such as making inferences, testing hypotheses and recognizing assumptions or they may refer to types of thinking such as conditional and deductive reasoning. For example, Ennis (1962) lists twelve aspects of critical thinking, Kolesnik (1962) defines eight concepts of critical thinking, Raths, Wasserman and Wasserman's (1969) curriculum materials focuses on twelve general abilities. The assumption is often made that students who receive sufficient practice at identifying

problems, drawing conclusions or evaluating syllogisms will be able to apply what they have learned to a wide variety of problems and issues.

Berlak (1965) criticized the "general aspects" view with respect to instruction in critical thinking. Berlak questions the value of teaching a set of generalized strategies. According to Berlak, the value of intellectual skills resides in their applicability to specific problems and issues of interest. Berlak recommends that investigators interested in teaching critical thinking, turn their attention to the thought processes and operations involved in specific areas of human experience. Aspects of critical thinking selected as behavioral objectives should "have demonstrated value to persons who have dealt successfully with some problem or issue." (p. 3).

The intent of the present investigation was to evaluate the effectiveness of a set of instructional materials designed to teach selected critical thinking skills to middle-school-age children. Lessons, games and activities were developed to form five independent courses or units. The units introduce standards for making judgments — specific principles, conventions and criteria for assessing the reliability, relevancy, warranty or sufficiency of information — within the context of life-like problem solving ventures. Each course is constructed around a distinct societal decision-making role and problem-solving paradigm:

1. "Advertising" — the role of the consumer advocate in evaluating persuasive techniques and empirical product claims; 2. "Judging" — the role of the courtroom lawyer in evaluating evidence and testimony;
3. "Reporting" — the role of the investigative reporter in interpreting reports and opinions; 4. "Conflict" — the role of the mediator in re-

solving interpersonal quarrels; and 5. "Causation" — the role of the social scientist in testing hypotheses and designing research studies.

Subjects

Six schools from the suburban Philadelphia area participated in the study. The socio-economic composition of these suburban communities ranged from lower-middle to upper-middle class. The building principal within each school suggested teachers who might be willing to take part in the study. For the most part, teachers who expressed a willingness to cooperate were randomly assigned to either the experimental or control condition. Then, students within each grade were randomly assigned to one of the two conditions. At the onset of the study, the groups were fairly equal in number within each grade. The total population consisted of 366 sixth-graders, 229 seventh-graders, 195 eighth-graders and 260 ninth-graders. By the end of the study, multiple absences and the withdrawal of one school reduced the total sample from 1,050 to 818 students.

Materials

The experimental material was the above-mentioned prototype curriculum package entitled *Making Judgments*. *Making Judgments* is comprised of five independent courses each consisting of five to six programmed lesson books and workbooks, one to two group board games and several teacher-led group activities. The units were constructed to provide a balance of individual small-group and large-group activities and the sequence of lessons within each unit was constructed to proceed from linearly programmed exercises to paper and pencil simulations or branched problem-solving tasks.

The criterion measures that were of principal interest in the study were the objective-referenced posttests that accompanied each course. Each of the five posttests consisted of approximately thirty items, usually two items for each objective. Some items were constructed in order to assess the mastery of concepts and principles while others were written in order to measure the extent to which students can transfer what they've learned to novel content areas. Transfer items were all open-ended reasoning items and did not include any of the course-related vocabulary present in many of the mastery items. Each course also included a pretest. Pretests were constructed by randomly selecting five mastery and five transfer items from the original pool of items.

Additional criterion measures included Level X of the Cornell Critical Thinking Test (Ennis and Millman, 1971) and a dispositional inventory made up of A Locus of Control Scale for Children (Nowicki and Strickland, 1972), Intolerance of Ambiguity Scale (Martin and Westie, 1959) and Attitude Toward School (Instructional Objectives Exchange, 1972).

All tests were administered with standard instructions and were scored by independent raters using standard scoring instructions. Protocols were coded to prevent the raters from knowing a student's name, school or experimental condition.

Formative evaluation questionnaires, interviews and observation scales were also used in the study to collect data concerning length and difficulty of the lesson, ease of administration of the materials, teacher and student attitude toward the material and teacher attitude toward the evaluation project.

Procedure

All cooperating teachers attended a brief orientation session in late September 1973. At that time, teachers were given an administration schedule and a Teacher's Guide for the first of the five units. From October to June, experimental teachers administered one unit after another according to a suggested schedule of two to three, one-hour sessions per week. Pretests were administered before and posttests were administered after each course for both the experimental and control groups. Control classes received no special treatment and were tested on the same day as the experimental class or classes within their school. Throughout the seven and one-half month study, a variety of formative evaluation indices were completed by teachers, students and observers. The Cornell Critical Thinking Test and the dispositional inventory were administered within a week of the post-test that accompanied the last of the five units.

Students who were absent from a session or two of any course were given makeup work to do at home or in a free period. Students who missed three or more lessons in a course were excluded from the analysis of that unit. The group sizes for the analyses are given in Table 1.

Table 1

SAMPLE SIZES FOR STATISTICAL ANALYSES

		TESTS						
Grade	Group	Unit A Posttest	Unit B Posttest	Unit C Posttest	Unit D Posttest	Unit E Posttest	Critical Thinking Test	Dispos. Inventory
6	E	198	202	203	205	205	204	
	C	146	149	156	151	151	147	
7	E	84	94	83	77	83	86	87
	C	78	106	89	89	83	84	88
8	E	71	63	49	56	55	49	52
	C	74	74	57	54	51	49	50
9	E	91	94	95	97	90	86	86
	C	92	95	85	86	90	84	79

The Analysis

A number of research questions were formulated including the following: 1) Are the treatment groups within each grade comparable? 2) What is happening that we need to consider for further modification of the materials? 3) How does sex, prior knowledge and intelligence affect the use of the materials? 4) Do the materials teach the content specified in the behavioral objectives (mastery)? 5) Do the materials teach the skills in a way that allows the student to generalize what he has learned to other situations (transfer)? 6) Do the students exhibit any differences in their disposition to use the skills taught? 7) What are the side effects which accompany the use of these materials?

Cochran's homogeneity of variance analysis was conducted on student IQ scores obtained from school records and on each of the five pretests for all grades and sites.

Responses to the dispositional inventory and the Cornell Critical Thinking Test were hand scored and subjected to analysis of variance (grade by treatment). The unit posttests were scored, then transferred to optical scanning sheets for a two-way analysis of variance conducted by computer.

Data from the formative evaluation instruments were summarized and are available elsewhere (Rose, 1974).

Results

All comparison groups were found to be homogeneous with respect to both IQ scores and pretest scores.

The experimental group outperformed the control group in 16 of the 20 comparisons made on the subtests of the Cornell Critical Thinking Test. In only one instance was a significant difference found between the groups' scores. Ninth grade experimental students scored significantly higher than ninth grade controls on the identifications of assumptions subtest.

Insert Table 2 here.

An analysis of variance performed on the subtests of the dispositional inventory showed a significant difference favoring experimental over control subjects on the Locus of Control measure in grade 6, on the Intolerance of Ambiguity measure in grade 7 and on the Attitude Toward School measure, also in grade 7. Differences favoring the control group over the experimental group were significant on the Attitude Toward School measure in grade 6 and on the Locus of Control measure in grade 9.

Insert Table 3 here.

Table 4 summarizes the results of the analysis of variance conducted on mastery items, transfer items and on total posttest scores for each course.

Insert Table 4 here.

Figure 1 displays the total posttest mean differences between experimental and control subjects by grade for all five courses. For the first course tested, Advertising, all comparisons are statistically significant and favor the experimental students in grades six, eight and nine.

TABLE 2

Means and Standard Deviations on Four Subtests of
the Cornell Critical Thinking Test
for Making Judgments and Control Groups

Grade	Site	N		Means								Standard Deviations							
				N=23 Hypothesis		N=23 Reliability		N=13 Premises		N=4 Assumptions		Hypothesis		Reliability		Premises		Assumptions	
		Exper.	Control	Exper.	Control	Exper.	Control	Exper.	Control	Exper.	Control	Exper.	Control	Exper.	Control	Exper.	Control	Exper.	Control
6	1	126	94	12.90	13.24	9.44	9.18	6.52	6.20	3.24	2.95	4.11	3.78	2.76	2.66	3.24	2.96	1.86	2.08
	2	78	53	13.41	13.40	9.35	9.38	7.05	6.21	3.85	3.32	4.29	4.83	2.96	2.47	3.40	3.33	1.93	2.36
7		86	84	13.90	13.99	10.21	10.32	8.70	8.68	4.41	4.18	4.68	3.48	3.39	3.31	3.73	3.43	2.37	2.08
8		49	49	13.90	14.39	10.18	10.69	8.31	8.08	4.08	3.94	3.45	3.72	3.35	3.63	3.47	3.68	2.11	2.53
9		86	84	13.69	13.04	10.49	10.04	7.40	6.68	4.30**	3.36	4.22	4.72	3.25	3.25	3.60	2.90	2.04	1.74

** $p < .01$.

TABLE 3

Means and Standard Deviations on Three Subtests
of Dispositional Inventory
for Making Judgments and Control Groups

Grade	Site	N		Means						Standard Deviations					
				Test A N=18		Test B N=40		Test C N=21		Test A N=18		Test B N=40		Test C N=21	
		Exper.	Control	Exper.	Control	Exper.	Control	Exper.	Control	Exper.	Control	Exper.	Control	Exper.	Control
6	1	125	97	6.91	6.22*	24.35	24.48	14.53	15.16*	2.32	2.07	4.70	4.72	3.55	2.91
	2	81	54	6.30	6.50	26.54	25.78	14.10	14.98	2.57	1.84	4.85	4.37	3.32	2.23
7		87	88	5.86	6.07	26.71*	25.28	14.93*	13.67	2.15	1.92	4.46	4.60	2.54	3.19
8		52	50	6.21	6.32	27.13	26.69	12.98	13.46	2.27	1.86	5.00	4.64	3.80	3.22
9		86	79	5.61	6.48*	28.14	26.85	12.10	12.44	2.50	2.45	5.41	4.45	5.62	3.40

Test A= Locus of Control
Test B= Intolerance of Ambiguity
Test C= Attitude toward School

* $p < .05$.

TABLE 4

SUMMARY OF F-RATIOS FOR
ANALYSIS OF VARIANCE OF POSTTEST SCORES
BY GRADE AND COURSE

COURSE	SUBTEST	GRADE 6	GRADE 7	GRADE 8	GRADE 9
Advertising	Mastery (28[a])	30.32**	5.41*	17.48**	45.86**
	Transfer (23)	17.50**	1	14.00**	52.06**
	Total Posttest	29.34**	3.14	18.55**	58.25**
Judging	Mastery (29)	43.74**	43.75**	20.20**	51.17**
	Transfer (24)	16.76**	11.37**	2.40	43.30**
	Total Posttest	36.31**	32.70**	11.1[illegible]**	52.51**
Reporting	Mastery (48)	22.19**	1.10	3.44	1.46
	Transfer (24)	4.53*	1	5.60*	2.05
	Total Posttest	20.08**	1.15	4.41*	1.70
Conflict	Mastery (70)	4.79*	2.71	4.35*	66.49**
	Transfer (7)	1	9.43*	3.41	31.51**
	Total Posttest	4.19*	3.68	4.55*	67.19**
Causation	Mastery (31)	2.68	1.07	4.18*	15.33**
	Transfer (20)	1.70	1.95	1	12.86**
	Total Posttest	2.55	1	2.80	15.89**

a Number of items in subtest.

* p .05.

** p .01.

Differences with respect to mastery items only were significant in grade seven. The results of the second course tested, Judging, showed significant differences favoring the treated students in all comparisons except for transfer items in grade eight. For Reporting, the third course tested, differences for both subtests and total tests were found to be significant in the sixth grade sample. In addition, the results for transfer items and total posttest items significantly favored the experimental group in grade eight.

Results for the Conflict unit showed significant differences favoring the experimental students on both subtests and total test in grade nine, on mastery items and total test in grades six and eight and on transfer items in grade seven. With respect to the final unit tested, Causation, again all comparisons significantly favored experimental students in grade nine, however, only one other comparison was significant. In grade eight, experimental students significantly outperformed controls on mastery items.

Insert Figure 1 here.

Discussion

The data lend some support to the premise that middle-school-age children can benefit from instructional materials designed to teach critical thinking skills in the context of problem solving and decision making ventures.

The results of the Cornell Critical Thinking Test and the subtests of the dispositional inventory can only be described as equivocal in the first instance and uninterpretable in the second. Insofar as these tests were

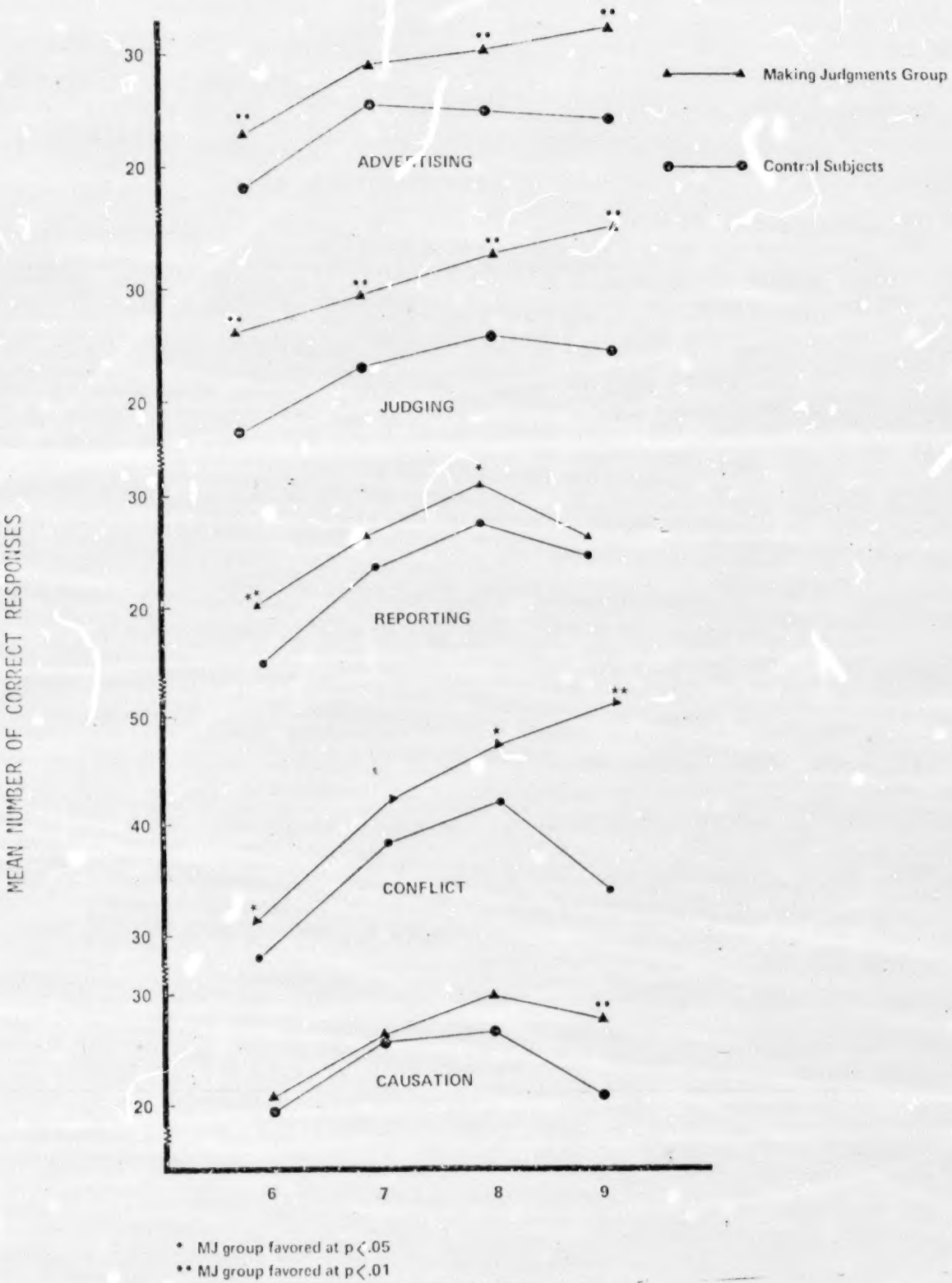

FIGURE I

ANALYSIS OF VARIANCE AS RELATED TO POSTTEST MEANS FOR TOTAL POSTTEST SCORES BY GRADE LEVEL FOR EACH COURSE

neither designed for this age group nor matched to the objectives of the instructional material, the lack of significant outcome was not regarded as a qualification of the materials' effectiveness.

The criterion measures of principal interest in this study were the mastery and transfer subtests of the objective-referenced posttests. Whereas the results of the first two courses tested strongly supported the effectiveness of the materials for accomplishing the objectives and fostering transfer of training, results from the final three units were less positive than expected.

Results from attitude questionnaires completed periodically by the teachers showed an increasing adverse reaction to the requirements and time involved in the pilot test that was unrelated to teachers' consistent positive evaluation of the materials. Printing and delivery delays increased the time span of the pilot test from four and one-half months to seven and one-half months. The decline in teacher morale precipitated by these delays may explain, in part, the decline in the effectiveness of units administered late in the study. In addition, the difficulty level of the last three units, Causation and Conflict in particular, was judged by teachers and others to be somewhat higher than the initial units. The fact that only ninth grade students outperformed controls on both subtests in the last two units lends support to this observation.

In retrospect, this study would have been more meaningful if all courses had been tested at the same time with independent samples. Such a field test of the package is currently being conducted in Los Angeles, Reno, Nevada, Philadelphia and West Germany. The five courses are being administered

as mini-courses with an equal amount of individual and teacher-led activities. To date, student and teacher attitudes toward the materials and the test conditions have been highly positive.

REFERENCES

Attitude Toward School, K-12, Los Angeles: Instructional Objectives Exchange, 1972.

Berlak, H. The teaching of thinking. *The School Review*, 1965, 73, 1-13.

Ennis, R. H. A concept of critical thinking. *Harvard Educational Review*, 1962, 32, 81-111.

Ennis, R. H., & Millman, J. *Cornell Critical Thinking Test*, Urbana, Illinois: University of Illinois Critical Thinking Project, 1971.

Kolesnik, W. B. Nature and improvement of students' thinking. *Catholic Educational Review*, 1964, 62, 231-240.

Martin, J. G., & Westie, F. R. The tolerant personality. *American Sociological Review*, 1959, 24, 521-528.

Nowicki, S., Jr., & Strickland, B. R. Locus of control scale for children, *Journal of Consulting and Clinical Psychology*, 1973, 40, 148-154.

Raths, L., Wasserman, J, & Wasserman, S. *The Thinking Box*, Westchester, Ill.: Benefic Press, 1969.

Rose, S. L. *Making Judgments: Pilot Test Report*. Unpublished working paper) Philadelphia: Research for Better Schools, 1974.

Russell, D. H. Research on the processes of thinking with some applications to reading. In R. G. Stauffer (Ed.). *Language and the Higher Thought Processes*. (A research bulletin prepared by a committee of the National Conference on Research in English.) Champaign, Ill.: National Council of Teachers of English, 1965.